SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Jung, Rudolf
              Beach, Larry R.
              Dress, Virginia M.
              Rao, A. Gururaj
              Ranch, Jerome P.
              Ertl, David S.
              Higgins, Regina K.

(ii) TITLE OF THE INVENTION: Alteration of Amino Acid Compositions in Seeds

(iii) NUMBER OF SEQUENCES: 13

(iv) CORRESPONDENCE ADDRESS:
  (A) ADDRESSEE: Pioneer Hi-Bred International, Inc.
  (B) STREET: 7100 NW 62nd Avenue, P.O. Box 1000
  (C) CITY: Johnston
  (D) STATE: IA
  (E) COUNTRY: USA
  (F) ZIP: 50131

(v) COMPUTER READABLE FORM:
  (A) MEDIUM TYPE: Diskette
  (B) COMPUTER: IBM Compatible
  (C) OPERATING SYSTEM: DOS
  (D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
  (A) APPLICATION NUMBER:
  (B) FILING DATE:
  (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
  (A) APPLICATION NUMBER:
  (B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
  (A) NAME: Michel, Marianne H
  (B) REGISTRATION NUMBER: 35,286
  (C) REFERENCE/DOCKET NUMBER: 0815

(ix) TELECOMMUNICATION INFORMATION:
  (A) TELEPHONE: 515-334-4467
  (B) TELEFAX: 515-334-6883
  (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 3363 base pairs
  (B) TYPE: nucleic acid

```
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: Other

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCGACCTCGA GGGGGGGCCC GGTACCCAGC TTTTGTTCCC TTTAGTGAGG GTTAATTGCG    60
CGCTTGGCGT AATCATGGTC ATAGCTGTTT CCTGTGTGAA ATTGTTATCC GCTCACAATT   120
CCACACAACA TACGAGCCGG AAGCATAAAG TGTAAAGCCT GGGGTGCCTA ATGAGTGAGC   180
TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTTCC AGTCGGGAAA CCTGTCGTGC   240
CAGCTGCATT AATGAATCGG CCAACGCGCG GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT   300
TCCGCTTCCT CGCTCACTGA CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG AGCGGTATCA   360
GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC   420
ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT   480
TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC GACGCTCAAG TCAGAGGTGG   540
CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC   600
TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC   660
GTGGCGCTTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC   720
AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC   780
TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC CACTGGCAGC AGCCACTGGT   840
AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG AGTTCTTGAA GTGGTGGCCT   900
AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC   960
TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT  1020
TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG GATCTCAAGA AGATCCTTTG  1080
ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT CACGTTAAGG GATTTTGGTC  1140
ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA ATTAAAAATG AAGTTTTAAA  1200
TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG  1260
GCACCTATCT CAGCGATCTG TCTATTTCGT TCATCCATAG TTGCCTGACT CCCCGTCGTG  1320
TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA  1380
GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC AGCCAGCCGG AAGGGCCGAG  1440
CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT CTATTAATTG TTGCCGGGAA  1500
GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC  1560
ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA  1620
AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCTCCG  1680
ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA TGGTTATGGC AGCACTGCAT  1740
AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG TGACTGGTGA GTACTCAACC  1800
AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT CTTGCCCGGC GTCAATACGG  1860
GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG  1920
GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT  1980
GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG TTTCTGGGTG AGCAAAAACA  2040
GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC GGAAATGTTG AATACTCATA  2100
CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC  2160
ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC CGCGCACATT TCCCCGAAAA  2220
GTGCCACCTA AATTGTAAGC GTTAATATTT TGTTAAAATT CGCGTTAAAT TTTTGTTAAA  2280
TCAGCTCATT TTTTAACCAA TAGGCCGAAA TCGGCAAAAT CCCTTATAAA TCAAAAGAAT  2340
AGACCGAGAT AGGGTTGAGT GTTGTTCCAG TTTGGAACAA GAGTCCACTA TTAAAGAACG  2400
TGGACTCCAA CGTCAAAGGG CGAAAAACCG TCTATCAGGG CGATGGCCCA CTACGTGAAC  2460
CATCACCCTA ATCAAGTTTT TTGGGGTCGA GGTGCCGTAA AGCACTAAAT CGGAACCCTA  2520
AAGGGAGCCC CCGATTTAGA GCTTGACGGG GAAAGCCGGC GAACGTGGCG AGAAAGGAAG  2580
GGAAGAAAGC GAAAGGAGCG GGCGCTAGGG CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG  2640
TAACCACCAC ACCCGCCGCG CTTAATGCGC CGCTACAGGG CGCGTCCCAT TCGCCATTCA  2700
GGCTGCGCAA CTGTTGGGAA GGGCGATCGG TGCGGGCCTC TTCGCTATTA CGCCAGCTGG  2760
CGAAAGGGGG ATGTGCTGCA AGGCGATTAA GTTGGGTAAC GCCAGGGTTT TCCCAGTCAC  2820
GACGTTGTAA AACGACGGCC AGTGAGCGCG CGTAATACGA CTCACTATAG GGCGAATTGG  2880
AGCTCCACCG CGGTGGCGGC CGCTCTAGAA CTAGTGGATC CGTCGACTAG AGGGCCCGAC  2940
GTCGAACTTA GGCACTAAGG GATGTGAGGC CAGCATCACC GTTGCAGAAA TTGACACAAG  3000
CATCACCACA ATTTTCCAAA TAGAGTTTCA TTTCTTCGTC GTCAGCAGCT GCGTTGACCA  3060
```

-42-

```
TGTAGTCACA CATGGAAGCC CTACACCCCA AGTTGCAATA CTTGACGGTG TCTGGTTCAT   3120
CTGAGTTGGA CACAAGGGCC AATTTGGGGA AGCCTGTAGG GCATTTTCCG CTACTTGTGA   3180
GTTTACACCT ACAGACGCCT GCGCATAACT TCTGAGCACC ACGGACGCGG CAAAGGTTGT   3240
AGCAGTTTCT TCCTAGGGTG CTCCTGCAGC AACTCTTGCC TTCTACTTGC ACCTGTTCGA   3300
GAACCAACCC CAGTATAAGT AAACACACCA TCACACCCTT GAGGCCCTTG CTGGTGGCCA   3360
TGG                                                                 3363
```
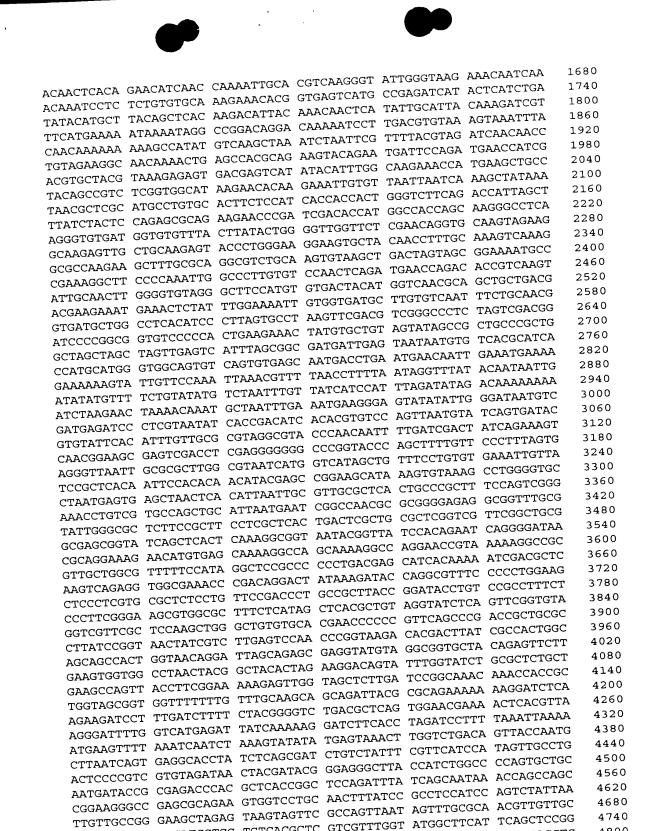
          (2) INFORMATION FOR SEQ ID NO:2:

          (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 3365 base pairs
              (B) TYPE: nucleic acid
              (C) STRANDEDNESS: single
              (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: Other

          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
TCGACCTCGA GGGGGGGCCC GGTACCCAGC TTTTGTTCCC TTTAGTGAGG GTTAATTGCG     60
CGCTTGGCGT AATCATGGTC ATAGCTGTTT CCTGTGTGAA ATTGTTATCC GCTCACAATT    120
CCACACAACA TACGAGCCGG AAGCATAAAG TGTAAAGCCT GGGGTGCCTA ATGAGTGAGC    180
TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTTCC AGTCGGGAAA CCTGTCGTGC    240
CAGCTGCATT AATGAATCGG CCAACGCGCG GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT    300
TCCGCTTCCT CGCTCACTGA CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG AGCGGTATCA    360
GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC    420
ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT    480
TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC GACGCTCAAG TCAGAGGTGG    540
CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC    600
TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC    660
GTGGCGCTTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC    720
AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC    780
TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC CACTGGCAGC AGCCACTGGT    840
AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG AGTTCTTGAA GTGGTGGCCT    900
AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC    960
TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT   1020
TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG GATCTCAAGA AGATCCTTTG   1080
ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT CACGTTAAGG GATTTTGGTC   1140
ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA ATTAAAAATG AAGTTTTAAA   1200
TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG   1260
GCACCTATCT CAGCGATCTG TCTATTTCGT TCATCCATAG TTGCCTGACT CCCCGTCGTG   1320
TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA   1380
GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC AGCCAGCCGG AAGGGCCGAG   1440
CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT CTATTAATTG TTGCCGGGAA   1500
GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC   1560
ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA   1620
AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCTCCG   1680
ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA TGGTTATGGC AGCACTGCAT   1740
AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG TGACTGGTGA GTACTCAACC   1800
AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT CTTGCCCGGC GTCAATACGG   1860
GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG   1920
GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT   1980
GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG TTTCTGGGTG AGCAAAAACA   2040
GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC GGAAATGTTG AATACTCATA   2100
CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC   2160
ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC CGCGCACATT TCCCCGAAAA   2220
GTGCCACCTA AATTGTAAGC GTTAATATTT TGTTAAAATT CGCGTTAAAT TTTTGTTAAA   2280
TCAGCTCATT TTTTAACCAA TAGGCCGAAA TCGGCAAAAT CCCTTATAAA TCAAAAGAAT   2340
```

                                    -43-

```
AGACCGAGAT AGGGTTGAGT GTTGTTCCAG TTTGGAACAA GAGTCCACTA TTAAAGAACG   2400
TGGACTCCAA CGTCAAAGGG CGAAAAACCG TCTATCAGGG CGATGGCCCA CTACGTGAAC   2460
CATCACCCTA ATCAAGTTTT TTGGGGTCGA GGTGCCGGC GAACGTGGCG AGAAAGGAAG   2520
AAGGGAGCCC CCGATTTAGA GCTTGACGGG GAAAGCCGGC GAACGTGGCG AGAAAGGAAG   2580
GGAAGAAAGC GAAAGGAGCG GGCGCTAGGG CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG   2640
TAACCACCAC ACCCGCCGCG CTTAATGCGC CGCTACAGGG CGCGTCCCAT TCGCCATTCA   2700
GGCTGCGCAA CTGTTGGGAA GGGCGATCGG TGCGGGCCTC TTCGCTATTA CGCCAGCTGG   2760
CGAAAGGGGG ATGTGCTGCA AGGCGATTAA GTTGGGTAAC GCCAGGGTTT TCCCAGTCAC   2820
GACGTTGTAA AACGACGGCC AGTGAGCGCG CGTAATACGA CTCACTATAG GGCGAATTGG   2880
AGCTCCACCG CGGTGGCGGC CGCTCTAGAA CTAGTGGATC CGTCGACTAG AGGGCCCGAC   2940
GTCGAACTTA GGCACTAAGG GATGTGAGGC CAGCATCACC GTTGCAGAAA TTGACACAAG   3000
CATCACCACA ATTTTCCAAA TAGAGTTTCA TTTCTTCGTC GTCAGCAGCT GCGTTGACCA   3060
TGTAGTCACA CATGGAAGCC CTACACCCCA AGTTGCAATA CTTGACGGTG TCTGGTTCAT   3120
CTGAGTTGGA CACAAGGGCC AATTTGGGGA AGCCTTTCGG GCATTTTCCG CTACTAGTCA   3180
GCTTACACTT GCAGACGCCT GCGCAAAGCT TCTTGGCGCC TTTGACTTTG CAAAGGTTGT   3240
AGCACTTCCT TCCCAGGGTA CTCTTGCAGC AACTCTTGCC TTCTACTTGC ACCTGTTCGA   3300
GAACCAACCC CAGTATAAGT AAACACACCA TCACACCCTT GAGGCCCTTG CTGGTGGCCA   3360
TGGTG                                                               3365
```

```
CTAAATTGTA AGCGTTAATA TTTTGTTAAA ATTCGCGTTA AATTTTTGTT AAATCAGCTC     60
ATTTTTTAAC CAATAGGCCG AAATCGGCAA AATCCCTTAT AAATCAAAAG AATAGACCGA    120
GATAGGGTTG AGTGTTGTTC CAGTTTGGAA CAAGAGTCCA CTATTAAAGA ACGTGGACTC    180
CAACGTCAAA GGGCGAAAAA CCGTCTATCA GGGCGATGGC CCACTACGTG AACCATCACC    240
CTAATCAAGT TTTTTGGGGT CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAAGGGAG    300
CCCCCGATTT AGAGCTTGAC GGGGAAAGCC GGCGAACGTG GCGAGAAAGG AAGGGAAGAA    360
AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG GTCACGCTGC GCGTAACCAC    420
CACACCCGCC GCGCTTAATG CGCCGCTACA GGGCGCGTCC CATTCGCCAT TCAGGCTGCG    480
CAACTGTTGG GAAGGGCGAT CGGTGCGGGC CTCTTCGCTA TTACGCCAGC TGGCGAAAGG    540
GGGATGTGCT GCAAGGCGAT TAAGTTGGGT AACGCCAGGG TTTTCCCAGT CACGACGTTG    600
TAAAACGACG GCCAGTGAGC GCGCGTAATA CGACTCACTA TAGGGCGAAT TGGAGCTCCA    660
CCGCGGTGGC GGCCGCTCTA GATTATATAA TTTATAAGCT AAACAACCCG GCCCTAAAGC    720
ACTATCGTAT CACCTATCTA AATAAGTCAC GGGAGTTTCG AACGTCCACT TCGTCGCACG    780
GAATTGCATG TTTCTTGTTG GAAGCATATT CACGCAATCT CCACACATAA AGGTTTATGT    840
ATAAACTTAC ATTTAGCTCA GTTTAATTAC AGTCTTATTT GGATGCATAT GTATGGTTCT    900
CAATCCATAT AAGTTAGAGT AAAAAATAAG TTTAAATTTT ATCTTAATTC ACTCCAACAT    960
ATATGGATCT ACAATACTCA TGTGCATCCA AACAAACTAC TTATATTGAG GTGAATTTGG   1020
TAGAAATTAA ACTAACTTAC ACACTAAGCC AATCTTTACT ATATTAAAGC ACCAGTTTCA   1080
ACGATCGTCC CGCGTCAATA TTATTAAAAA ACTCCTACAT TTCTTTATAA TCAACCCGCA   1140
CTCTTATAAT CTCTTCTCTA CTACTATAAT AAGAGAGTTT ATGTACAAAA TAAGGTGAAA   1200
TTATCTATAA GTGTTCTGGA TATTGGTTGT TGGCTCCCAT ATTCACACAA CCTAATCAAT   1260
AGAAACATA TGTTTTATTA AAACAAAATT TATCATATAT CATATATATA TATATATCAT   1320
ATATATATAT AAACCGTAGC AATGCACGGG CATATAACTA GTGCAACTTA ATACATGTGT   1380
GTATTAAGAT GAATAAGAGG GTATCCAAAT AAAAAACTTG TTGCTTACGT ATGGATCGAA   1440
AGGGGTTGGA AACGATTAAA CGATTAAATC TCTTCCTAGT CAAAATTGAA TAGAAGGAGA   1500
TTTAATATAT CCCAATCCCC TTCGATCATC CAGGTGCAAC CGTATAAGTC CTAAAGTGGT   1560
GAGGAACACG AAAGAACCAT GCATTGGCAT GTAAAGCTCC AAGAATTTGT TGTATCCTTA   1620
```

```
ACAACTCACA GAACATCAAC CAAAATTGCA CGTCAAGGGT ATTGGGTAAG AAACAATCAA    1680
ACAAATCCTC TCTGTGTGCA AAGAAACACG GTGAGTCATG CCGAGATCAT ACTCATCTGA    1740
TATACATGCT TACAGCTCAC AAGACATTAC AAACAACTCA TATTGCATTA CAAAGATCGT    1800
TTCATGAAAA ATAAAATAGG CCGGACAGGA CAAAAATCCT TGACGTGTAA AGTAAATTTA    1860
CAACAAAAAA AAAGCCATAT GTCAAGCTAA ATCTAATTCG TTTTACGTAG ATCAACAACC    1920
TGTAGAAGGC AACAAAACTG AGCCACGCAG AAGTACAGAA TGATTCCAGA TGAACCATCG    1980
ACGTGCTACG TAAAGAGAGT GACGAGTCAT ATACATTTGG CAAGAAACCA TGAAGCTGCC    2040
TACAGCCGTC TCGGTGGCAT AAGAACACAA GAAATTGTGT TAATTAATCA AAGCTATAAA    2100
TAACGCTCGC ATGCCTGTGC ACTTCTCCAT CACCACCACT GGGTCTTCAG ACCATTAGCT    2160
TTATCTACTC CAGAGCGCAG AAGAACCCGA TCGACACCAT GGCCACCAGG TG CAAGTAGAAG    2220
AGGGTGTGAT GGTGTGTTTA CTTATACTGG GGTTGGTTCT CGAACAGGTG CAAGTAGAAG    2280
GCAAGAGTTG CTGCAAGAGT ACCCTGGGAA GGAAGTGCTA CAACCTTTGC AAAGTCAAAG    2340
GCGCCAAGAA GCTTTGCGCA GGCGTCTGCA AGTGTAAGCT GACTAGTAGC GGAAAATGCC    2400
CGAAAGGCTT CCCCAAATTG GCCCTTGTGT CCAACTCAGA TGAACCAGAC ACCGTCAAGT    2460
ATTGCAACTT GGGGTGTAGG GCTTCCATGT GTGACTACAT GGTCAACGCA GCTGCTGACG    2520
ACGAAGAAAT GAAACTCTAT TTGGAAAATT GTGGTGATGC TTGTGTCAAT TTCTGCAACG    2580
GTGATGCTGG CCTCACATCC CTTAGTGCCT AAGTTCGACG TCGGGCCCTC TAGTCGACGG    2640
ATCCCCGGCG GTGTCCCCCA CTGAAGAAAC TATGTGCTGT AGTATAGCCG CTGCCCGCTG    2700
GCTAGCTAGC TAGTTGAGTC ATTTAGCGGC GATGATTGAG TAATAATGTG TCACGCATCA    2760
CCATGCATGG GTGGCAGTGT CAGTGTGAGC AATGACCTGA ATGAACAATT GAAATGAAAA    2820
GAAAAAAGTA TTGTTCCAAA TTAAACGTTT TAACCTTTTA ATAGGTTTAT ACAATAATTG    2880
ATATATGTTT TCTGTATATG TCTAATTTGT TATCATCCAT TTAGATATAG ACAAAAAAAA    2940
ATCTAAGAAC TAAAACAAAT GCTAATTTGA AATGAAGGGA GTATATATTG GGATAATGTC    3000
GATGAGATCC CTCGTAATAT CACCGACATC ACACGTGTCC AGTTAATGTA TCAGTGATAC    3060
GTGTATTCAC ATTTGTTGCG CGTAGGCGTA CCCAACAATT TTGATCGACT ATCAGAAAGT    3120
CAACGGAAGC GAGTCGACCT CGAGGGGGGG CCCGGTACCC AGCTTTTGTT CCCTTTAGTG    3180
AGGGTTAATT GCGCGCTTGG CGTAATCATG GTCATAGCTG TTTCCTGTGT GAAATTGTTA    3240
TCCGCTCACA ATTCCACACA ACATACGAGC CGGAAGCATA AAGTGTAAAG CCTGGGGTGC    3300
CTAATGAGTG AGCTAACTCA CATTAATTGC GTTGCGCTCA CTGCCCGCTT TCCAGTCGGG    3360
AAACCTGTCG TGCCAGCTGC ATTAATGAAT CGGCCAACGC GCGGGGAGAG GCGGTTTGCG    3420
TATTGGGCGC TCTTCCGCTT CCTCGCTCAC TGACTCGCTG CGCTCGGTCG TTCGGCTGCG    3480
GCGAGCGGTA TCAGCTCACT CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA    3540
CGCAGGAAAG AACATGTGAG CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC    3600
GTTGCTGGCG TTTTTCCATA GGCTCCGCCC CCTGACGAG CATCACAAAA ATCGACGCTC    3660
AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC CAGGCGTTTC CCCCTGGAAG    3720
CTCCCTCGTG CGCTCTCCTG TTCCGACCCT GCCGCTTACC GGATACCTGT CCGCCTTTCT    3780
CCCTTCGGGA AGCGTGGCGC TTTCTCATAG CTCACGCTGT AGGTATCTCA GTTCGGTGTA    3840
GGTCGTTCGC TCCAAGCTGG GCTGTGTGCA CGAACCCCCC GTTCAGCCCG ACCGCTGCGC    3900
CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA CACGACTTAT CGCCACTGGC    3960
AGCAGCCACT GGTAACAGGA TTAGCAGAGC GAGGTATGTA GGCGGTGCTA CAGAGTTCTT    4020
GAAGTGGTGG CCTAACTACG GCTACACTAG AAGGACAGTA TTTGGTATCT GCGCTCTGCT    4080
GAAGCCAGTT ACCTTCGGAA AAAGAGTTGG TAGCTCTTGA TCCGGCAAAC AAACCACCGC    4140
TGGTAGCGGT GGTTTTTTTG TTTGCAAGCA GCAGATTACG CGCAGAAAAA AAGGATCTCA    4200
AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG TGGAACGAAA ACTCACGTTA    4260
AGGGATTTTG GTCATGAGAT TATCAAAAAG GATCTTCACC TAGATCCTTT TAAATTAAAA    4320
ATGAAGTTTT AAATCAATCT AAAGTATATA TGAGTAAACT TGGTCTGACA GTTACCAATG    4380
CTTAATCAGT GAGGCACCTA TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG    4440
ACTCCCCGTC GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGCTGC    4500
AATGATACCG CGAGACCCAC GCTCACCGGC TCCAGATTTA TCAGCAATAA ACCAGCCAGC    4560
CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC AACTTTATCC GCCTCCATCC AGTCTATTAA    4620
TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC    4680
CATTGCTACA GGCATCGTGG TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG    4740
TTCCCAACGA TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC    4800
CTTCGGTCCT CCGATCGTTG TCAGAAGTAA GTTGGCCGCA GTGTTATCAC TCATGGTTAT    4860
GGCAGCACTG CATAATTCTC TTACTGTCAT GCCATCCGTA AGATGCTTTT CTGTGACTGG    4920
TGAGTACTCA ACCAAGTCAT TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC    4980
GGCGTCAATA CGGGATAATA CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG    5040
AAAACGTTCT TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT    5100
```

```
GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG    5160
GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG    5220
TTGAATACTC ATACTCTTCC TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT    5280
CATGAGCGGA TACATATTTG AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC    5340
ATTTCCCCGA AAAGTGCCAC                                                5360
```
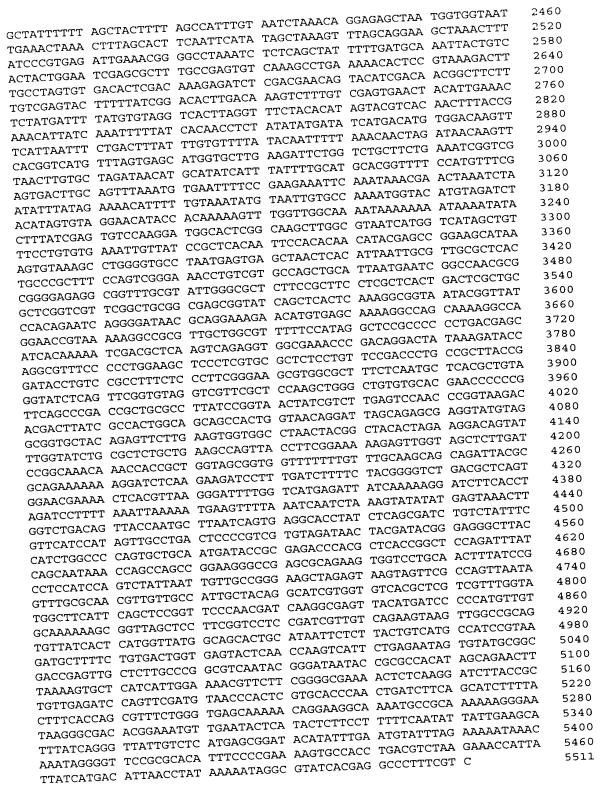
          (2) INFORMATION FOR SEQ ID NO:4:

          (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 5511 base pairs
               (B) TYPE: nucleic acid
               (C) STRANDEDNESS: single
               (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: Other
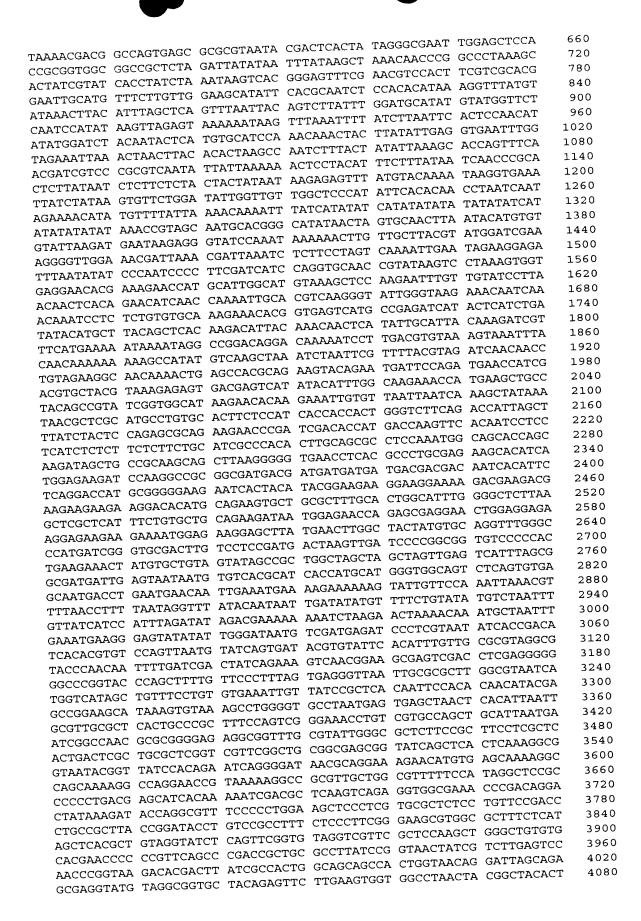
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA      60
CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG     120
TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC     180
ACCATATGCG GTGTGAAATA CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGGCGCC     240
ATTCGCCATT CAGGCTGCGC AACTGTTGGG AAGGGCGATC GGTGCGGGCC TCTTCGCTAT     300
TACGCCAGCT GGCGAAAGGG GGATGTGCTG CAAGGCGATT AAGTTGGGTA ACGCCAGGGT     360
TTTCCCAGTC ACGACGTTGT AAAACGACGG CCAGTGAATT CTTTTATGAA TAATAATAAT     420
GCATATCTGT GCATTACTAC CTGGGATACA AGGGCTTCTC CGCCATAACA AATTGAGTTG     480
CGATGCTGAG AACGAACGGG GAAGAAAGTA AGCGCCGCCC AAAAAAAACG AACATGTACG     540
TCGGCTATAG CAGGTGAAAG TTCGTGCGCC AATGAAAAGG GAACGATATG CGTTGGGTAG     600
TTGGGATACT TAAATTTGGA GAGTTTGTTG CATACACTAA TCCACTAAAG TTGTCTATCT     660
TTTTAACAGC TCTAGGCAGG ATATAAGATT TATATCTAAT CTGTTGGAGT TGCTTTTAGA     720
GTAACTTTTC TCTCTGTTTC GTTTATAGCC GATTAGCACA AAATTAAACT AGGTGACGAG     780
AAATAAAGAA AAACGGAGGC AGTAAAAAAT ACCCAAAAAA ATACTTGGAG ATTTTTGTCT     840
CAAAATTATC TTCTAATTTT AAAAGCTACA TATTAAAAAT ACTATATATT AAAAATACTT     900
CGAGATCATT GCTTGGGATG GGCAGGGCCA ATAGCTAATT GCTAAGGATG GGCTATATTT     960
ATGTATCGTC TGAAACATGT AGGGGCTAAT AGTTAGATGA CTAATTTGCT GTGTTCGTAC    1020
GGGGTGCTGT TTGAGCCTAG CGATGAAGGG TCATAGTTTC ATACAAGAAC TCACTTTTGG    1080
TTCGTCTGCT GTGTCTGTTC TCAGCGTAAC GGCATCAATG GATGCCAAAC TCCGCAAGGG    1140
GACAAATGAA GAAGCGAAGA GATTATAGAA CACGCACGTG TCATTATTTA TTTATGGACT    1200
TGCCTCAGTA GCTTACAGCA TCGTACCCGC ACGTACATAC TACAGAGCCA CACTTATTGC    1260
ACTGCCTGCC GCTTACGTAC ATAGTTAACA CGCAGAGAGG TATATACATA CACGTCCAAC    1320
GTCTCCACTC AGGCTCATGC TACGTACGCA CGTCGGTCGC GCGCCACCCT CTCGTTGCTT    1380
CCTGCTCGTT TTGGCGAGCT AGAGGGCCCG ACGTCGAACT TAGGCACTAA GGGATGTGAG    1440
GCCAGCATCA CCGTTGCAGA AATTGACACA AGCATCACCA CAATTTTCCA AATAGAGTTT    1500
CATTTCTTCG TCGTCAGCAG CTGCGTTGAC CATGTAGTCA CACATGGAAG CCCTACACCC    1560
CAAGTTGCAA TACTTGACGG TGTCTGGTTC ATCTGAGTTG GACACAAGGG CCAATTTGGG    1620
GAAGCCTTTC GGGCATTTTC CGCTACTAGT CAGCTTACAC TTGCAGACGC CTGCGCAAAG    1680
CTTCTTGGCG CCTTTGACTT TGCAAAGGTT GTAGCACTTC CTTCCCAGGG TACTCTTGCA    1740
GCAACTCTTG CCTTCTACTT GCACCTGTTC GAGAACCAAC CCCAGTATAA GTAAACACAC    1800
CATCACACCC TTGAGGCCCT TGCTGGTGGC CATGGTGTAG TGTCGACTGT GATATCCTCG    1860
GGTGTGTGTT GGATCCTTGG GTTGGCTGTA TGCAGAACTA AAGCGGAGGT GGCGCGCATT    1920
TATACCAGCG CCGGGCCCTG GTACGTGGCG CGGCCGCGCG GCTACGTGGA GGAAGGCTGC    1980
GTGGCAGCAG ACACACGGGT CGCCACGTCC CGCCGTACTC TCCTTACCGT GCTTATCCGG    2040
GCTCCGGCTC GGTGCACGCC AGGGTGTGGC CGCCTCTGAG CAGACTTTGT CGTGTTCCAC    2100
AGTGGTGTCG TGTTCCGGGG ACTCCGATCC GCGGCGAGCG ACCGAGCGTG TAAAAGAGTT    2160
CCTACTAGGT ACGTTCATTG TATCTGGACG ACGGGCAGCG GACAATTTGC TGTAAGAGAG    2220
GGGCAGTTTT TTTTTAGAAA AACAGAGAAT TCCGTTGAGC TAATTGTAAT TCAACAAATA    2280
AGCTATTAGT TGGTTTTAGC TTAGATTAAA GAAGCTAACG ACTAATAGCT AATAATTAGT    2340
TGGTCTATTA GTTGACTCAT TTTAAGGCCC TGTTTCAATC TCGCGAGATA AACTTTAGCA    2400
```

                                    -46-

```
GCTATTTTTT AGCTACTTTT AGCCATTTGT AATCTAAACA GGAGAGCTAA TGGTGGTAAT    2460
TGAAACTAAA CTTTAGCACT TCAATTCATA TAGCTAAAGT TTAGCAGGAA GCTAAACTTT    2520
ATCCCGTGAG ATTGAAACGG GGCCTAAATC TCTCAGCTAT TTTTGATGCA AATTACTGTC    2580
ACTACTGGAA TCGAGCGCTT TGCCGAGTGT CAAAGCCTGA AAAACACTCC GTAAAGACTT    2640
TGCCTAGTGT GACACTCGAC AAAGAGATCT CGACGAACAG TACATCGACA ACGGCTTCTT    2700
TGTCGAGTAC TTTTTATCGG ACACTTGACA AAGTCTTTGT CGAGTGAACT ACATTGAAAC    2760
TCTATGATTT TATGTGTAGG TCACTTAGGT TTCTACACAT AGTACGTCAC AACTTTACCG    2820
AAACATTATC AAATTTTTAT CACAACCTCT ATATATGATA TCATGACATG TGGACAAGTT    2880
TCATTAATTT CTGACTTTAT TTGTGTTTTA TACAATTTTT AAACAACTAG ATAACAAGTT    2940
CACGGTCATG TTTAGTGAGC ATGGTGCTTG AAGATTCTGG TCTGCTTCTG AAATCGGTCG    3000
TAACTTGTGC TAGATAACAT GCATATCATT TATTTTGCAT GCACGGTTTT CCATGTTTCG    3060
AGTGACTTGC AGTTTAAATG TGAATTTTCC GAAGAAATTC AAATAAACGA ACTAAATCTA    3120
ATATTTATAG AAAACATTTT TGTAAATATG TAATTGTGCC AAAATGGTAC ATGTAGATCT    3180
ACATAGTGTA GGAACATACC ACAAAAAGTT TGGTTGGCAA AATAAAAAAA ATAAAATATA    3240
CTTTATCGAG TGTCCAAGGA TGGCACTCGG CAAGCTTGGC GTAATCATGG TCATAGCTGT    3300
TTCCTGTGTG AAATTGTTAT CCGCTCACAA TTCCACACAA CATACGAGCC GGAAGCATAA    3360
AGTGTAAAGC CTGGGGTGCC TAATGAGTGA GCTAACTCAC ATTAATTGCG TTGCGCTCAC    3420
TGCCCGCTTT CCAGTCGGGA AACCTGTCGT GCCAGCTGCA TTAATGAATC GGCCAACGCG    3480
CGGGGAGAGG CGGTTTGCGT ATTGGGCGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC    3540
GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA ATACGGTTAT    3600
CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA    3660
GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC    3720
ATCACAAAAA TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC    3780
AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG    3840
GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC TCACGCTGTA    3900
GGTATCTCAG TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG    3960
TTCAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC    4020
ACGACTTATC GCCACTGGCA GCAGCCACTG GTAACAGGAT TAGCAGAGCG AGGTATGTAG    4080
GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACTACGG CTACACTAGA AGGACAGTAT    4140
TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT    4200
CCGGCAAACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC    4260
GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT GACGCTCAGT    4320
GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ATCTTCACCT    4380
AGATCCTTTT AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT    4440
GGTCTGACAG TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC    4500
GTTCATCCAT AGTTGCCTGA CTCCCCGTCG TGTAGATAAC TACGATACGG GAGGGCTTAC    4560
CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT CCAGATTTAT    4620
CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG    4680
CCTCCATCCA GTCTATTAAT TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA    4740
GTTTGCGCAA CGTTGTTGCC ATTGCTACAG GCATCGTGGT GTCACGCTCG TCGTTTGGTA    4800
TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC CCCATGTTGT    4860
GCAAAAAAGC GGTTAGCTCC TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG    4920
TGTTATCACT CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCCGTAA    4980
GATGCTTTTC TGTGACTGGT GAGTACTCAA CCAAGTCATT CTGAGAATAG TGTATGCGGC    5040
GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT AGCAGAACTT    5100
TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC    5160
TGTTGAGATC CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA    5220
CTTTCACCAG CGTTTCTGGG TGAGCAAAAA CAGGAAGGCA AAATGCCGCA AAAAAGGGAA    5280
TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT TATTGAAGCA    5340
TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC    5400
AAATAGGGGT TCCGCGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA GAAACCATTA    5460
TTATCATGAC ATTAACCTAT AAAAATAGGC GTATCACGAG GCCCTTTCGT C             5511
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 5115 base pairs
    (B) TYPE: nucleic acid

-47-

```
       (C)  STRANDEDNESS: single
       (D)  TOPOLOGY: linear

    (ii)  MOLECULE TYPE: Other

    (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTTGGGAGCT CTCCCATATG GTCGACCTGC AGGCGGCCGC TCTAGAACTA GTGGATCCCC      60
CCCTCGAGGT CGACGGTATC GATAAGCTTG ATATCTTACA AGGCCCAGCC CAGCGACCTA     120
TTACACAGCC CGCTCGGGCC CGCGACGTCG GGACACATCT TCTTCCCCCT TTTGGTGAAG     180
CTCTGCTCGC AGCTGTCCGG CTCCTTGGAC GTTCGTGTGG CAGATTCATC TGTTGTCTCG     240
TCTCCTGTGC TTCCTGGGTA GCTTGTGTAG TGGAGCTGAC ATGGTCTGAG CAGGCTTAAA     300
ATTTGCTCGT AGACGAGGAG TACCAGCACA GCACGTTGCG GATTTCTCTG CCTGTGAAGT     360
GCAACGTCTA GGATTGTCAC ACGCCTTGGT CGCGTCGCGT CGCGTCGCGT CGATGCGGTG     420
GTGAGCAGAG CAGCAACAGC TGGGCGGCCC AACGTTGGCT TCCGTGTCTT CGTCGTACGT     480
ACGCGCGCGC CGGGGACACG CAGCAGAGAG CGGAGAGCGA GCCGTGCACG GGGAGGTGGT     540
GTGGAAGTGG AGCCGCGCGC CCGGCCGCCC GCGCCGGTG GGCAACCCAA AAGTACCCAC     600
GACAAGCGAA GGCGCCAAAG CGATCCAAGC TCCGGAACGC AACAGCATGC GTCGCGTCGG     660
AGAGCCAGCC ACAAGCAGCC GAGAACCGAA CCGGTGGGCG ACGCGTCATG GGACGGACGC     720
GGGCGACGCT TCCAAACGGG CCACGTACGC CGGCGTGTGC GTGCGTGCAG ACGACAAGCC     780
AAGGCGAGGC AGCCCCCGAT CGGGGAACGG GTATCGTGGG GGGCGCGGGC GGAGGAGAGC     840
GTCGCTGGTG CGCAGTGCCG GGGGGAACGG GTATCGTGGG GGGCGCGGGC GGAGGAGAGC     900
GTGGCGAGGG CCGAGAGCAG CGCGCGGCCG GGTCACGCAA CGCGCCCCAC GTACTGCCCT     960
CCCCCTCCGC GCGCGCTAGA AATACCGAGG CCTGGACCGG GGGGGGGCCC CGTCACATCC    1020
ATCCATCGAC CGATCGATCG CCACAGCCAA CACCACCCGC CGAGGCGACG CGACAGCCGC    1080
CAGGAGGAAG GAATAAACTC ACTGCCAGCC AGTGAAGGGG GAGAAGTGTA CTGCTCCGTC    1140
GACCAGTGCG CGCACCGCCC GGCAGGGCTG CTCATCTCGT CGACGACCAG GTTCTGTTCC    1200
GATCCGATCC GATCCTGTCC TTGAGTTTCG TCCAGATCCT GGCGCGTATC TGCGTGTTTG    1260
ATGATCCAGG TTCTTCGAAC CTAAATCTGT CCGTGCACAC GTCTTTTCTC TCTCTCCTAC    1320
GCAGTGGATT AATCGCCATG GCCACCAGCA AGGGCCTCAA GGGTGTGATG GTGTGTTTAC    1380
TTATACTGGG GTTGGTTCTC GAACAGGTGC AAGTAGAAGG CAAGAGTTGC TGCAAGAGTA    1440
CCCTGGGAAG GAAGTGCTAC AACCTTTGCA AAGTCAAAGG CGCCAAGAAG CTTTGCGCAG    1500
GCGTCTGCAA GTGTAAGCTG ACTAGTAGCG GAAAATGCCC GAAAGGCTTC CCCAAATTGG    1560
CCCTTGTGTC CAACTCAGAT GAACCAGACA CCGTCAAGTA TTGCAACTTG GGGTGTAGGG    1620
CTTCCATGTG TGACTACATG GTCAACGCAG CTGCTGACGA CGAAGAAATG AAACTCTATT    1680
TGGAAAATTG TGGTGATGCT TGTGTCAATT TCTGCAACGG TGATGCTGGC CTCACATCCC    1740
TTAGTGCCTA AGTTCGACGT CGGGCCCTCT AGATGCGGCC CGGGTGAAGA GTTCGCCCTG    1800
CAGGGCCCCT GATCTCGCGC GTGGTGCAAA GATGTTGGGA CATCTTCTTA TATATGCTGT    1860
TTCGCTTATG TGATATGGAC AAGTATGTGT AGATGCTTGC TTGTGCTAGT GTAATGTAGT    1920
GTAGTGGTGG CCAGTGGCAC AACCTAATAA GCGCATGAAC TAATTGCTTG CGTGTGTAGT    1980
TAAGTACCGA TCGGTAATTT TATATTGCGA GTAAATAAAT GGACCTGTAG TGGTGGAGTA    2040
AATAATCCCT GCTGTTCGGT GTTCTTATCG CTCCTCGTAT AGATATTATA TAGAGTACAT    2100
TTTTCTCTCT CTGAATCCTA CGTGTGTGAA ATTTCTATAT CATTACTGTA AAATTTCTGC    2160
GTTCCAAAAG AGACCATAGC CTATCTTTGG CCCTGTTTGT TTCGGCTTCT GGCAGCTTCT    2220
GGCCACCAAA AGCTGCTGCG GACTGCCAAA CGCTCAGATT TTCAGCTAGC TTCTATAAAA    2280
TTAGTTGGGG CAAAAACCAT CCAAAATCAA TATAAACACA TAATCGGTTG AGTCGTTGTA    2340
ATATTAGGAA TCTGTCACTT TCTAGATCCT GAGCCCTATG AACAACTTTA TCTTTCTCCA    2400
TACGTAATCG TAATGATACT CAGATTCTCT CCACAGCCAG ATTCTCCTCA CAGCCAGATT    2460
TTCAGAAAAG CTGGTCAGAA AAAAGTTAAA CCAAACAGAC CCTTTGTGTA TGCATGGATC    2520
GGCTTTCCCC GTCAAGCTCT AAATCGGGGG CTCCCTTTAG GGTTCCGATT TAGAGCTTTA    2580
CGGCACCTCG ACCGCAAAAA ACTTGATTTG GGTGATGGTT CACGTAGTGG GCCATCGCCC    2640
TGATAGACGG TTTTTCGCCC TTTGACGTTG GAGTCCACGT TCTTTAATAG TGGACTCTTG    2700
TTCCAAACTG GAACAACACT CAACCCTATC TCGGTCTATT CTTTTGATTT ATAAGGGATT    2760
TTGCCGATTT CGGCCTATTG GTTAAAAAAT GAGCTGATTT AACAAATATT TAACGCGAAT    2820
TTTAACAAAA TATTAACGTT TACAATTTCG CCTGATGCGG TATTTTCTCC TTACGCATCT    2880
GTGCGGTATT TCACACCGCA TACAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT    2940
ATTTGTTTAT TTTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA    3000
TAAATGCTTC AATAATATTG AAAAAGGAAG AGTATGAGTA TTCAACATTT CCGTGTCGCC    3060
```

                              -48-

```
CTTATTCCCT TTTTTGCGGC ATTTTGCCTT CCTGTTTTTG CTCACCCAGA AACGCTGGTG   3120
AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC   3180
AACAGCGGTA AGATCCTTGA GAGTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT   3240
TTTAAAGTTC TGCTATGTCA TACACTATTA TCCCGTATTG ACGCCGGGCA AGAGCAACTC   3300
GGTCGCCGGG CGCGGTATTC TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG   3360
CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT   3420
AACACTGCGG CCAACTTACT TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCGCTTTT   3480
TTGCACAACA TGGGGGATCA TGTAACTCGC CTTGATCGTT GGGAACCGGA GCTGAATGAA   3540
GCCATACCAA ACGACGAGCG TGACACCACG ATGCCTGTAG CAATGCCAAC AACGTTGCGC   3600
AAACTATTAA CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAATTAAT AGACTGGATG   3660
GAGGCGGATA AAGTTGCAGG ACCACTTCTG CGCTCGGCCC TTCCGGCTGG CTGGTTTATT   3720
GCTGATAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGGTA TCATTGCAGC ACTGGGGCCA   3780
GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT   3840
GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAACTGTCA   3900
GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAAC TTCATTTTTA ATTTAAAAGG   3960
ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTTCG   4020
TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT   4080
CTGCGCGTAA TCTGCTGCTT GCAAACAAAA AAACCACCGC TACCAGCGGT GGTTTGTTTG   4140
CCGGATCAAG AGCTACCAAC TCTTTTTCCG AAGGTAACTG GCTTCAGCAG AGCGCAGATA   4200
CCAAATACTG TCCTTCTAGT GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA   4260
CCGCCTACAT ACCTCGCTCT GCTAATCCTG TTACCAGTGG CTGCTGCCAG TGGCGATAAG   4320
TCGTGTCTTA CCGGGTTGGA CTCAAGACGA TAGTTACCGG ATAAGGCGCA GCGGTCGGGC   4380
TGAACGGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA   4440
TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GGCGGACAGG   4500
TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGGAAAC   4560
GCCTGGTATC TTTATAGTCC TGTCGGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTTG   4620
TGATGCTCGT CAGGGGGGCG GAGCCTATCG AAAAACGCCA GCAACGCGGC CTTTTTACGG   4680
TTCCTGGCCT TTTGCTGGCC TTTTGCTCAC ATGTTCTTTC CTGCGTTATC CCCTGATTCT   4740
GTGGATAACC GTATTACCGC CTTTGAGTGA GCTGATACCG CTCGCCGCAG CCGAACGACC   4800
GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC CAATACGCAA ACCGCCTCTC   4860
CCCGCGCGTT GGCCGATTCA TTAATGCAGC TGGCACGACA GGTTTCCCGA CTGGAAAGCG   4920
GGCAGTGAGC GCAACGCAAT TAATGTGAGT TAGCTCACTC ATTAGGCACC CCAGGCTTTA   4980
CACTTTATGC TTCCGGCTCG TATGTTGTGT GGAATTGTGA GCGGATAACA ATTTCACACA   5040
GGAAACAGCT ATGACCATGA TTACGCCAAG CTATTTAGGT GACACTATAG AATACTCAAG   5100
CTATGCATCC AACGC                                                   5115
```

        (2) INFORMATION FOR SEQ ID NO:6:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 5392 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: Other

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
CTAAATTGTA AGCGTTAATA TTTTGTTAAA ATTCGCGTTA AATTTTTGTT AAATCAGCTC    60
ATTTTTTAAC CAATAGGCCG AAATCGGCAA AATCCCTTAT AAATCAAAAG AATAGACCGA   120
GATAGGGTTG AGTGTTGTTC CAGTTTGGAA CAAGAGTCCA CTATTAAAGA ACGTGGACTC   180
CAACGTCAAA GGGCGAAAAA CCGTCTATCA GGGCGATGGC CCACTACGTG AACCATCACC   240
CTAATCAAGT TTTTTGGGGT CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAAGGGAG   300
CCCCCGATTT AGAGCTTGAC GGGGAAAGCC GGCGAACGTG GCGAGAAAGG AAGGGAAGAA   360
AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG GTCACGCTGC GCGTAACCAC   420
CACACCCGCC GCGCTTAATG CGCCGCTACA GGGCGCGTCC CATTCGCCAT TCAGGCTGCG   480
CAACTGTTGG GAAGGGCGAT CGGTGCGGGC CTCTTCGCTA TTACGCCAGC TGGCGAAAGG   540
GGGATGTGCT GCAAGGCGAT TAAGTTGGGT AACGCCAGGG TTTTCCCAGT CACGACGTTG   600
```

                                  -49-

```
TAAAACGACG GCCAGTGAGC GCGCGTAATA CGACTCACTA TAGGGCGAAT TGGAGCTCCA  660
CCGCGGTGGC GGCCGCTCTA GATTATATAA TTTATAAGCT AAACAACCCG GCCCTAAAGC  720
ACTATCGTAT CACCTATCTA AATAAGTCAC GGGAGTTTCG AACGTCCACT TCGTCGCACG  780
GAATTGCATG TTTCTTGTTG GAAGCATATT CACGCAATCT CCACACATAA AGGTTTATGT  840
ATAAACTTAC ATTTAGCTCA GTTTAATTAC AGTCTTATTT GGATGCATAT GTATGGTTCT  900
CAATCCATAT AAGTTAGAGT AAAAAATAAG TTTAAATTTT ATCTTAATTC ACTCCAACAT  960
ATATGGATCT ACAATACTCA TGTGCATCCA AACAAACTAC TTATATTGAG GTGAATTTGG  1020
TAGAAATTAA ACTAACTTAC ACACTAAGCC AATCTTTACT ATATTAAAGC ACCAGTTTCA  1080
ACGATCGTCC CGCGTCAATA TTATTAAAAA ACTCCTACAT TTCTTTATAA TCAACCCGCA  1140
CTCTTATAAT CTCTTCTCTA CTACTATAAT AAGAGAGTTT ATGTACAAAA TAAGGTGAAA  1200
TTATCTATAA GTGTTCTGGA TATTGGTTGT TGGCTCCCAT ATTCACATATA TATATATCAT  1260
AGAAACATA TGTTTTATTA AAACAAAATT TATCATATAT CATATATATA TATATATCAT  1320
ATATATATAT AAACCGTAGC AATGCACGGG CATATAACTA GTGCAACTTA ATACATGTGT  1380
GTATTAAGAT GAATAAGAGG GTATCCAAAT AAAAAACTTG TTGCTTACGT ATGGATCGAA  1440
AGGGGTTGGA AACGATTAAA CGATTAAATC TCTTCCTAGT CAAAATTGAA TAGAAGGAGA  1500
TTTAATATAT CCCAATCCCC TTCGATCATC CAGGTGCAAC CGTATAAGTC CTAAAGTGGT  1560
GAGGAACACG AAAGAACCAT GCATTGGCAT GTAAAGCTCC AAGAATTTGT TGTATCCTTA  1620
ACAACTCACA GAACATCAAC CAAAATTGCA CGTCAAGGGT ATTGGGTAAG AAACAATCAA  1680
ACAAATCCTC TCTGTGTGCA AAGAAACACG GTGAGTCATG CCGAGATCAT ACTCATCTGA  1740
TATACATGCT TACAGCTCAC AAGACATTAC AAACAACTCA TATTGCATTA CAAAGATCGT  1800
TTCATGAAAA ATAAAATAGG CCGGACAGGA CAAAAATCCT TGACGTGTAA AGTAAATTTA  1860
CAACAAAAAA AAAGCCATAT GTCAAGCTAA ATCTAATTCG TTTTACGTAG ATCAACAACC  1920
TGTAGAAGGC AACAAAACTG AGCCACGCAG AAGTACAGAA TGATTCCAGA TGAACCATCG  1980
ACGTGCTACG TAAAGAGAGT GACGAGTCAT ATACATTTGG CAAGAAACCA TGAAGCTGCC  2040
TACAGCCGTA TCGGTGGCAT AAGAACACAA GAAATTGTGT TAATTAATCA AAGCTATAAA  2100
TAACGCTCGC ATGCCTGTGC ACTTCTCCAT CACCACCACT GGGTCTTCAG ACCATTAGCT  2160
TTATCTACTC CAGAGCGCAG AAGAACCCGA TCGACACCAT GACCAAGTTC ACAATCCTCC  2220
TCATCTCTCT TCTCTTCTGC ATCGCCCACA CTTGCAGCGC CTCCAAATGG CAGCACCAGC  2280
AAGATAGCTG CCGCAAGCAG CTTAAGGGGG TGAACCTCAC GCCCTGCGAG AAGCACATCA  2340
TGGAGAAGAT CCAAGGCCGC GGCGATGACG ATGATGATGA TGACGACGAC AATCACATTC  2400
TCAGGACCAT GCGGGGGAAG AATCACTACA TACGGAAGAA GGAAGGAAAA GACGAAGACG  2460
AAGAAGAAGA AGGACACATG CAGAAGTGCT GCGCTTTGCA CTGGCATTTG GGGCTCTTAA  2520
GCTCGCTCAT TTCTGTGCTG CAGAAGATAA TGGAGAACCA GAGCGAGGAA CTGGAGGAGA  2580
AGGAGAAGAA GAAAATGGAG AAGGAGCTTA TGAACTTGGC TACTATGTGC AGGTTTGGGC  2640
CCATGATCGG GTGCGACTTG TCCTCCGATG ACTAAGTTGA TCCCCGGCGG TGTCCCCCAC  2700
TGAAGAAACT ATGTGCTGTA GTATAGCCGC TGGCTAGCTA GCTAGTTGAG TCATTTAGCG  2760
GCGATGATTG AGTAATAATG TGTCACGCAT CACCATGCAT GGGTGGCAGT CTCAGTGTGA  2820
GCAATGACCT GAATGAACAA TTGAAATGAA AAGAAAAAAG TATTGTTCCA AATTAAACGT  2880
TTTAACCTTT TAATAGGTTT ATACAATAAT TGATATATGT TTTCTGTATA TGTCTAATTT  2940
GTTATCATCC ATTTAGATAT AGACGAAAAA AAATCTAAGA ACTAAACAA ATGCTAATTT  3000
GAAATGAAGG GAGTATATAT TGGGATAATG TCGATGAGAT CCCTCGTAAT ATCACCGACA  3060
TCACACGTGT CCAGTTAATG TATCAGTGAT ACGTGTATTC ACATTGTTG CGCGTAGGCG  3120
TACCCAACAA TTTTGATCGA CTATCAGAAA GTCAACGGAA GCGAGTCGAC CTCGAGGGGG  3180
GGCCCGGTAC CCAGCTTTTG TTCCCTTTAG TGAGGGTTAA TTGCGCGCTT GGCGTAATCA  3240
TGGTCATAGC TGTTTCCTGT GTGAAATTGT TATCCGCTCA CAATTCCACA CAACATACGA  3300
GCCGGAAGCA TAAAGTGTAA AGCCTGGGGT GCCTAATGAG TGAGCTAACT CACATTAATT  3360
GCGTTGCGCT CACTGCCCGC TTTCCAGTCG GGAAACCTGT CGTGCCAGCT GCATTAATGA  3420
ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTTCCGC TTCCTCGCTC  3480
ACTGACTCGC TGCGCTCGGT CGTTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG  3540
GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC  3600
CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC  3660
CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA  3720
CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC  3780
CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAT  3840
AGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG  3900
CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC  3960
AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA  4020
GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT  4080
```

```
AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT   4140
GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG   4200
CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG   4260
TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA   4320
AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA   4380
TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG   4440
ATCTGTCTAT TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA   4500
CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG   4560
GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAAGCTAG AGTAAGTAGT   4620
TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC   4680
TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA   4740
TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT   4800
AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC   4860
ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA   4920
TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGCGCCA   4980
CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG AAAACTCTCA   5040
AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT   5100
TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC   5160
GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTTCAA   5220
TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT   5280
TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC AC           5392
```

(Note: the line numbering in the right column as printed reads: 4140, 4200, 4260, 4320, 4380, 4440, 4500, 4560, 4620, 4680, 4740, 4800, 4860, 4920, 4980, 5040, 5100, 5160, 5220, 5280, 5340, 5392.)

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 5173 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: Other

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
CTAAATTGTA AGCGTTAATA TTTTGTTAAA ATTCGCGTTA AATTTTTGTT AAATCAGCTC     60
ATTTTTTAAC CAATAGGCCG AAATCGGCAA AATCCCTTAT AAATCAAAAG AATAGACCGA    120
GATAGGGTTG AGTGTTGTTC CAGTTTGGAA CAAGAGTCCA CTATTAAAGA ACGTGGACTC    180
CAACGTCAAA GGGCGAAAAA CCGTCTATCA GGGCGATGGC CCACTACGTG AACCATCACC    240
CTAATCAAGT TTTTTGGGGT CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAAGGGAG    300
CCCCCGATTT AGAGCTTGAC GGGGAAAGCC GGCGAACGTG GCGAGAAAGG AAGGGAAGAA    360
AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG GTCACGCTGC GCGTAACCAC    420
CACACCCGCC GCGCTTAATG CGCCGCTACA GGGCGCGTCC CATTCGCCAT TCAGGCTGCG    480
CAACTGTTGG GAAGGGCGAT CGGTGCGGGC CTCTTCGCTA TTACGCCAGC TGGCGAAAGG    540
GGGATGTGCT GCAAGGCGAT TAAGTTGGGT AACGCCAGGG TTTTCCCAGT CACGACGTTG    600
TAAAACGACG GCCAGTGAGC GCGCGTAATA CGACTCACTA TAGGGCGAAT TGGAGCTCCA    660
CCGCGGTGGC GGCCGCTCTA GATTATATAA TTTATAAGCT AAACAACCCG GCCCTAAAGC    720
ACTATCGTAT CACCTATCTA AATAAGTCAC GGGAGTTTCG AACGTCCACT TCGTCGCACG    780
GAATTGCATG TTTCTTGTTG GAAGCATATT CACGCAATCT CCACACATAA AGGTTTATGT    840
ATAAACTTAC ATTTAGCTCA GTTTAATTAC AGTCTTATTT GGATGCATAT GTATGGTTCT    900
CAATCCATAT AAGTTAGAGT AAAAAATAAG TTTAAATTTT ATCTTAATTC ACTCCAACAT    960
ATATGGATCT ACAATACTCA TGTGCATCCA AACAAACTAC TTATATTGAG GTGAATTTGG   1020
TAGAAATTAA ACTAACTTAC ACACTAAGCC AATCTTTACT ATATTAAAGC ACCAGTTTCA   1080
ACGATCGTCC CGCGTCAATA TTATTAAAAA ACTCCTACAT TTCTTTATAA TCAACCCGCA   1140
CTCTTATAAT CTCTTCTCTA CTACTATAAT AAGAGAGTTT ATGTACAAAA TAAGGTGAAA   1200
TTATCTATAA GTGTTCTGGA TATTGGTTGT TGGCTCCCAT ATTCACACAA CCTAATCAAT   1260
AGAAAACATA TGTTTTATTA AAACAAAATT TATCATATAT CATATATATA TATATATCAT   1320
ATATATATAT AAACCGTAGC AATGCACGGG CATATAACTA GTGCAACTTA ATACATGTGT   1380
```

```
GTATTAAGAT GAATAAGAGG GTATCCAAAT AAAAAACTTG TTGCTTACGT ATGGATCGAA   1440
AGGGGTTGGA AACGATTAAA CGATTAAATC TCTTCCTAGT CAAAATTGAA TAGAAGGAGA   1500
TTTAATATAT CCCAATCCCC TTCGATCATC CAGGTGCAAC CGTATAAGTC CTAAAGTGGT   1560
GAGGAACACG AAAGAACCAT GCATTGGCAT GTAAAGCTCC AAGAATTTGT TGTATCCTTA   1620
ACAACTCACA GAACATCAAC CAAAATTGCA CGTCAAGGGT ATTGGGTAAG AAACAATCAA   1680
ACAAATCCTC TCTGTGTGCA AAGAAACACG GTGAGTCATG CCGAGATCAT ACTCATCTGA   1740
TATACATGCT TACAGCTCAC AAGACATTAC AAACAACTCA TATTGCATTA CAAAGATCGT   1800
TTCATGAAAA ATAAAATAGG CCGGACAGGA CAAAAATCCT TGACGTGTAA AGTAAATTTA   1860
CAACAAAAAA AAAGCCATAT GTCAAGCTAA ATCTAATTCG TTTTACGTAG ATCAACAACC   1920
TGTAGAAGGC AACAAAACTG AGCCACGCAG AAGTACAGAA TGATTCCAGA TGAACCATCG   1980
ACGTGCTACG TAAAGAGAGT GACGAGTCAT ATACATTTGG CAAGAAACCA TGAAGCTGCC   2040
TACAGCCGTA TCGGTGGCAT AAGAACACAA GAATTGTGT TAATTAATCA AAGCTATAAA   2100
TAACGCTCGC ATGCCTGTGC ACTTCTCCAT CACCACCACT GGGTCTTCAG ACCATTAGCT   2160
TTATCTACTC CAGAGCGCAG AAGAACCCGA TCGACACCAT GAAGTCGGTG GAGAAGAAAC   2220
CGAAGGGTGT GAAGACAGGT GCGGGTGACA AGCATAAGCT GAAGACAGAG TGGCCGGAGT   2280
TGGTGGGGAA ATCGGTGGAG AAAGCCAAGA AGGTGATCCT GAAGGACAAG CCAGAGGCGC   2340
AAATCATAGT TCTACCGGTT GGTACAAAGG TGGGTAAGCA TTATAAGATC GACAAGGTCA   2400
AGCTTTTTGT GGATAAAAAG GACAACATCG CGCAGGTCCC CAGGGTCGGC TAGCCTCGAG   2460
ATCCCCGGCG GTGTCCCCCA CTGAAGAAAC TATGTGCTGT AGTATAGCCG CTGGCTAGCT   2520
AGCTAGTTGA GTCATTTAGC GGCGATGATT GAGTAATAAT GTGTCACGCA TCACCATGCA   2580
TGGGTGGCAG TCTCAGTGTG AGCAATGACC TGAATGAACA ATTGAAATGA AAAGAAAAAA   2640
GTATTGTTCC AAATTAAACG TTTTAACCTT TTAATAGGTT TATACAATAA TTGATATATG   2700
TTTTCTGTAT ATGTCTAATT TGTTATCATC CATTTAGATA TAGACGAAAA AAAATCTAAG   2760
AACTAAAACA AATGCTAATT TGAAATGAAG GGAGTATATA TTGGGATAAT GTCGATGAGA   2820
TCCCTCGTAA TATCACCGAC ATCACACGTG TCCAGTTAAT GTATCAGTGA TACGTGTATT   2880
CACATTTGTT GCGCGTAGGC GTACCCAACA ATTTTGATCG ACTATCAGAA AGTCAACGGA   2940
AGCGAGTCGA CCTCGAGGGG GGGCCCGGTA CCCAGCTTTT GTTCCCTTTA GTGAGGGTTA   3000
ATTGCGCGCT TGGCGTAATC ATGGTCATAG CTGTTTCCTG TGTGAAATTG TTATCCGCTC   3060
ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA   3120
GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACTGCCCG CTTTCCAGTC GGGAAACCTG   3180
TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG   3240
CGCTCTTCCG CTTCCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCGGCT GCGGCGAGCG   3300
GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA TAACGCAGGA   3360
AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG   3420
GCGTTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG   3480
AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC   3540
GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT TCTCCCTTCG   3600
GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT GTAGGTCGTT   3660
CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTCAGC CCGACCGCTG CGCCTTATCC   3720
GGTAACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC   3780
ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG   3840
TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA   3900
GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC   3960
GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC TCAAGAAGAT   4020
CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG AAAACTCACG TTAAGGGATT   4080
TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT   4140
TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC   4200
AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTCAT CCATAGTTGC CTGACTCCCC   4260
GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA   4320
CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG   4380
GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC   4440
CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT   4500
ACAGGCATCG TGGTGTCACG CTCGTCGTTT GGTATGGCTT CATTCAGCTC CGGTTCCCAA   4560
CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT   4620
CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTTAT CACTCATGGT TATGGCAGCA   4680
CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC   4740
TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGCGACCGA GTTGCTCTTG CCCGGCGTCA   4800
ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT   4860
```

-52-

```
TCTTCGGGGC GAAAACTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC    4920
ACTCGTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA    4980
AAAACAGGAA GGCAAAATGC CGCAAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA    5040
CTCATACTCT TCCTTTTTCA ATATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC    5100
GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTTCCC    5160
                                                                     5173
CGAAAAGTGC CAC
```

    (2) INFORMATION FOR SEQ ID NO:8:

      (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 54 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: Other

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
AGTATAAGTA AACACACCAT CACACCCTTG AGGCCCTTGC TGGTGGCCAT GGTG          54
```

        (2) INFORMATION FOR SEQ ID NO:9:

      (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 55 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: Other

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
CCTCACATCC CTTAGTGCCT AAGTTCGACG TCGGGCCCTC TAGTCGACGG ATCCA         55
```

        (2) INFORMATION FOR SEQ ID NO:10:

      (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 35 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: Other

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
AGCGGAAAAT GCCCGAAAGG CTTCCCCAAA TTGGC                               35
```

        (2) INFORMATION FOR SEQ ID NO:11:

      (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 45 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGCGCAGGCG TCTGCAAGTG TAAGCTGACT AGTAGCGGAA AATGC   45

        (2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 50 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: Other

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TACAACCTTT GCAAAGTCAA AGGCGCCAAG AAGCTTTGCG CAGGCGTCTG   50

        (2) INFORMATION FOR SEQ ID NO:13:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 50 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: Other

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GCAAGAGTTG CTGCAAGAGT ACCCTGGGAA GGAAGTGCTA CAACCTTTGC   50

The invention is not limited to the exact details shown and described, for it should be understood that many variations and modifications may be made while remaining within the spirit and scope of the invention defined by the claims.